SEMANTIC MAPPING OF STEM CONTENTS FOR AURAL REPRESENTATION USING LITERATURE MINING

Venkatesh Bharadwaj (Mathew Palakal[1], Steven Mannheimer[1]),
Department of Computer and Information Science, School of Science,
Indiana University–Purdue University Indianapolis, Indianapolis, IN 46202.

As STEM education increasingly relies on illustrations, animations and video to communicate complex concepts, blind and visually impaired (BVI) students are increasingly left behind. However, tablet computers and other digital technologies offer the potential for a sound-based solution that leverages the ability of BVI students to "think aurally" beyond simple spoken terminology. Previous work has shown that non-verbal sound can improve educational outcomes for BVI students. The challenge is translating science concepts that may be essentially soundless (e.g. photosynthesis or cumulus clouds) into sounds that communicate the component ideas of a concept. One key is to consider any science concept as a *process* or *activity* with actions and actors, and to identify sounds that refer to them. Our research focuses on computational strategies for analyzing the sentences used in standard K-12 textbooks to define or describe any given science concept-activity, and generate a semantic sequence of words which correlates to sounds that can best portray or embody them. This is done with the help of Natural Language Processing (NLP) tools in combination with a newly developed Information Extraction (IE) algorithm. Because each word in a semantic sequence can potentially correlate to multiple sounds, it is necessary to find a dynamic path connecting the list of sounds that represent a word sequence in the context of the given science process or categorical domain. For example, there are multiple sounds associated with the basic concept "water:" e.g. *splashing*, *pouring*, drops *dripping*. But in the context of "precipitation" *dripping* is most relevant. The algorithm to identify the best concept-to-sound correlations is a newly developed, self-learning and adaptive algorithm. This research supports, and is informed by, experiments in aural pedagogy conducted at Indiana School of Blind and Visually Impaired. Our long-term goal is the generation of a language of non-verbal sounds.

[1]School of Informatics, Indiana University-Purdue University, Indianapolis, IN-46202

Funding from Google Research and National Science Foundation.